Generalized least squares estimation of linear models containing rational future expectations

Nijman, T.E.; Palm, F.C.

Published in:
International Economic Review

Publication date:
1991

INTERNATIONAL ECONOMIC REVIEW
Vol. 32, No. 2, May 1991

GENERALIZED LEAST SQUARES ESTIMATION OF LINEAR MODELS CONTAINING RATIONAL FUTURE EXPECTATIONS*

BY THEO NIJMAN AND FRANZ PALM[1]

We discuss the choice of approximations for unobserved expectations underlying consistent estimators in linear RE models with future expectations. We show how estimators which are more efficient than the commonly used GMM estimators can be obtained if it is assumed that the future expectation depends on a finite number of variables only. Numerical results for a simple model illustrate the relative efficiency of various estimators.

1. INTRODUCTION

Models with expectational variables are widely used in empirical econometric research. Various estimators have been put forward to estimate the parameters in these models (see e.g. the surveys by Pesaran 1988 and Nijman 1990). Many of these estimators belong to the general class of Generalized Method of Moments (GMM) estimators proposed by Hansen (1982). If a GMM estimator is used, the unobserved expectations are approximated by the corresponding realization following a suggestion of McCallum (1976). For static models an alternative referred to as the substitution approach by Wickens (1982) consists in fitting an auxiliary regression and approximating the unobserved expectation by the projection from the auxiliary equation. In this note we show how to approximate the future expectation by the projection from an auxiliary regression and obtain a generalized least squares estimator (GLS) that is at least as efficient as the GMM estimator based on future realizations as proxies for the future expectations, provided one is willing to assume that the future expectation depends on a finite number of variables only. Related GLS estimators in models with current expectations have been considered by Hoffman (1987), Pesaran (1988, p. 168) and Nijman and Steel (1988). Note that Pesaran (1988) claims that the computation of the GLS estimator is complicated, but does not refer to the binomial inversion lemma used in the Appendix of this paper. Both the GLS estimator proposed in this paper and the GMM estimators are not fully efficient in general. Both estimators however are often computationally much more attractive than the efficient maximum likelihood (ML) estimator.

The paper is organized as follows. In Section 2 we introduce the GLS estimator and compare it with GMM estimators. In Section 3 numerical results on the relative efficiency of the various estimators illustrate the argument. Finally Section 4 contains some concluding remarks.

* Manuscript received January 1989; revised May 1990.

[1] The authors thank Michael McAleer, Neil Ericsson and three unknown referees for helpful comments.

2. THE GLS ESTIMATOR

Consider the following linear model for the scalar variable y_t

$$(1) \qquad y_t = \rho E[y_{t+1}|I_t] + \alpha y_{t-1} + \delta' x_t + e_t, \qquad e_t \sim IN(0, \sigma_e^2),$$

while the vector x_t is generated by

$$(2) \qquad x_t = \sum_{i=1}^{p} \Gamma_i x_{t-i} + v_t, \qquad v_t \sim IN(0, \Theta).$$

Assume that e_t and v_s are independent for all t and s, that v_t is independent of x_{t-1}, x_{t-2}, ... and define $I_t = \{y_t, x_t, y_{t-1}, x_{t-1}, ...\}$. Throughout this paper we assume that the model in (1) and (2) is identified, which implies e.g. $p \geq 2$ (see Pesaran 1988, section 6.3.4). Equation (1) describes a standard RE model with future expectation. Well-known examples are models of factor demand under adjustment costs and Cagan's money demand equation (see, e.g., Pesaran 1988 or Nijman 1990). In the latter example, $\alpha = 0$ a priori. Equation (2) states that the k-dimensional vector of exogenous variables is generated by a vector autoregression, possibly with restrictions on the parameter matrices Γ_i. If x_t and y_t are assumed to be stationary and if one of the roots of the characteristic equation $\rho z^2 - z + \alpha = 0$ falls inside and the other outside the unit circle, the model has a unique stationary solution which can be characterized by

$$(3) \qquad y_{t+1} = \lambda y_t + \sum_{i=0}^{p-1} {}_t\psi_i' x_{t-i} + u_t, \qquad (|\lambda| < 1)$$

where u_t is independent of I_t and the parameters λ and ψ_i are in general highly nonlinear function of the structural parameters in (1) and (2). Note that $\lambda = 0$, if there is no lagged endogenous variable in (1) as in the Cagan money demand equation which simplifies the analysis below somewhat. By construction, the error term u_t in (3) satisfies

$$(4) \qquad Eu_t u_{t+k} = 0 \text{if } k \neq 0; \quad Eu_t^2 = \sigma_u^2;$$

$$Eu_t e_{t+k} = 0 \quad \text{if } k \neq 1; \quad Eu_t e_{t+1} = \sigma_{eu}.$$

As (3) implies that $E[y_{t+1}|I_t] = \lambda y_t + \Sigma_{i=0}^{p-1} \psi_i' x_{t-i}$, ML estimation comes down to joint estimation of (2) and

$$(5) \qquad y_t = (1 - \rho\lambda)^{-1}\{\alpha y_{t-1} + (\rho\psi_0' + \delta')x_t + \rho \sum_{i=1}^{p-1} \psi_i' x_{t-i} + e_t\}$$

imposing all the restrictions on the ψ's which can be computationally very demanding.

The class of GMM estimators is based on substitution of (3) into (1), along the lines suggested by McCallum (1976), which yields

$$y_t = \rho y_{t+1} + \alpha y_{t-1} + \delta' x_t + e_t - \rho u_t \tag{6}$$

from which ρ and δ can be consistently estimated using instrumental variables (IV) methods because $x_t, y_{t-1}, x_{t-1}, \ldots$ are orthogonal to e_t and u_t. As $e_t - \rho u_t$ is autocorrelated (see (4)), the standard IV estimators can be improved upon by exploiting the properties of this error term as proposed by Cumby, Huizinga and Obstfeld (1983) and Hayashi and Sims (1983) for the present linear model and by Hansen (1982) for the general case. Note that all GMM estimators are based on the fact that all variables in I_t are orthogonal to $y_{t+1} - E[y_{t+1}|I_t]$ but do not use the restrictions on the parameters of (3).

An alternative class of estimators starts with an auxiliary equation derived from (3) relating y_{t+1} to the valid instruments,

$$y_{t+1} = \lambda^2 y_{t-1} + \psi_0' x_t + \sum_{i=1}^{p-1} (\psi_i' + \lambda \psi_{i-1}') x_{t-i} + \lambda \psi_{p-1}' x_{t-p} + u_t + \lambda u_{t-1}. \tag{7}$$

Zero restrictions on the matrices Γ_i in (2) will often yield exclusion restrictions on (7) which can easily be taken into account by writing

$$y_{t+1} = z_t' \pi + \xi_t, \tag{8}$$

where z_t is a vector consisting of all elements of $(y_{t-1}, x_t', \ldots, x_{t-p}')$ with nonzero coefficient in (7), π is a corresponding vector of parameters and $\xi_t = u_t + \lambda u_{t-1}$. Substituting $E[y_{t+1}|I_t] = z_t'\pi + \lambda u_{t-1}$ in (1) and replacing π by $\hat{\pi}$, the OLS estimator of π from (8), one obtains

$$y_t = \rho z_t' \hat{\pi} + \alpha y_{t-1} + \delta' x_t + \varepsilon_t + \rho z_t'(\pi - \hat{\pi}), \tag{9}$$

where $\varepsilon_t = e_t + \rho\lambda u_{t-1}$. Note that $z_t'\hat{\pi}$ is used to approximate $E[y_{t+1}|I_t]$ rather than $z_t'\hat{\pi} + \lambda u_{t-1}$ to obtain orthogonality between the regressors and the disturbances and note that z_t and u_{t-1} are uncorrelated because of the assumed independence of e_t and v_s. In obvious vector notation (9) reads as

$$y = W\vartheta + \eta, \tag{10}$$

with $W = [Z(Z'Z)^{-1}Z'y_+ : y_- : X]$, $\vartheta' = (\rho, \alpha, \delta')$ and $\eta = \varepsilon - \rho Z(Z'Z)^{-1}Z'\xi$.

If ρ, α and δ are estimated from (10) using OLS the resulting estimator of course coincides with the IV estimator from (6) when z_t is the vector of instruments because of the two stage least squares interpretation of IV estimators. The parameters in (10) however can be alternatively estimated by GLS which yields a more efficient estimator,

$$\hat{\vartheta}_{GLS} = (W'\hat{\Omega}^{-1}W)^{-1}W'\hat{\Omega}^{-1}y \tag{11}$$

provided a matrix $\hat{\Omega}^{-1}$ can be found such that

$$W'\hat{\Omega}^{-1}\eta/\sqrt{T} \overset{d}{\sim} N(0, \operatorname{plim} T^{-1}W'\hat{\Omega}^{-1}W), \tag{12}$$

where T denotes the sample size. In the Appendix we show that the inverse of the covariance matrix Ω of η conditional on Z satisfies (12) and present expressions for this matrix. Evidently this matrix depends on ρ, λ, σ_e^2, σ_{eu} and σ_u^2 but these parameters can be replaced by consistent estimates, without affecting the limiting distribution. Consistent estimates of λ and σ_u^2 can be obtained from OLS on (3), while OLS on (9) yields a consistent estimate of ρ. The residuals of these regressions as well as the estimates of ρ, λ and σ_u^2 can then be straightforwardly used to estimate σ_e^2 and σ_{eu} consistently. In the Appendix we also show that the GLS estimator can be interpreted as an IV estimator of (11) with a matrix weighted linear combination of W, Z and Z_- as the matrix of instruments, which proves its consistency. Note that GLS type estimators of (6) are inconsistent as noted e.g. by Cumby et al. (1983) or Hansen and Hodrick (1980) in a related problem.

An important reason for using the GLS estimator is that it is at least as efficient as any GMM estimator which is based on the orthogonality conditions in (6) only. This is true because the estimator proposed by Cumby, Huizinga and Obstfeld (1983) can be considered as an estimator based on pre-multiplication of (10) by Z' and therefore cannot be more efficient than GLS on (10). Moreover it is well-known (see e.g. Hansen 1982) that a GMM estimator based on (6) only, cannot be more efficient than the estimator proposed by Cumby et al. if $x_t, y_{t-1}, x_{t-1}, \cdots, y_{t-r}, x_{t-r}$ is used as the vector of instruments with r tending to infinity. The GLS estimator can of course be more efficient than GMM estimators because the zero restrictions in (3) are taken into account. A GMM estimator which simultaneously imposes the orthogonality restrictions in (3) and (6) in a bivariate model will probably be as efficient as the GLS estimator, but this estimator is no longer computationally attractive.

3. NUMERICAL EXAMPLE

In order to illustrate the argument in the previous section we will now present numerical results on the relative asymptotic efficiency of the various estimators for the very simple case where α is known to be zero, $p = 2$ and $k = 1$, with k being the number of exogenous variables in (1). It can easily be checked that in this case $\lambda = 0$,

$$\psi_0 = \delta(\gamma_1 + \rho\gamma_2)\{1 - \rho\gamma_1 - \rho^2\gamma_2\}^{-1} \tag{13}$$

$$\psi_1 = \delta\gamma_2\{1 - \rho\gamma_1 - \rho^2\gamma_2\}^{-1} \tag{14}$$

and

$$u_t = e_{t+1} + (\rho\psi_0 + \delta)v_{t+1}, \tag{15}$$

where lower case γ's indicate that we consider a scalar case. Using (6) it is straightforward to evaluate the asymptotic variances of the estimators. Moreover it can be shown that the relative efficiency does not depend on all six parameters in the model but depends on ρ, γ_1, γ_2 and $R^2 = E(y_t - e_t)^2/Ey_t^2$ only.

In Table 1 the asymptotic efficiency of four estimators of ρ and δ is presented.

TABLE 1
RELATIVE EFFICIENCY OF THE MAXIMUM LIKELIHOOD ESTIMATOR COMPARED WITH ALTERNATIVE ESTIMATORS FOR ρ AND δ IF $k = 1$, $p = 2$ AND $\rho = 0.9$

			Rel. Eff. $\hat{\rho}$				Rel. Eff. $\hat{\delta}$			
R^2	γ_1	γ_2	IV	CHO	HS	GLS	IV	CHO	HS	GLS
.5	1.2	−.35	1.87	1.45	1.32	1.19	2.01	1.53	1.37	1.22
.5	1.4	−.45	2.30	1.62	1.30	1.13	2.43	1.69	1.33	1.15
.5	1.5	−.56	2.29	1.63	1.34	1.17	2.45	1.71	1.39	1.19
.5	1.7	−.72	2.91	1.89	1.32	1.12	3.06	1.96	1.34	1.13
.9	1.2	−.35	1.23	1.18	1.18	1.17	1.23	1.18	1.17	1.16
.9	1.4	−.45	1.46	1.30	1.28	1.23	1.46	1.30	1.28	1.23
.9	1.5	−.56	1.35	1.24	1.22	1.20	1.33	1.22	1.21	1.19
.9	1.7	−.72	1.54	1.31	1.27	1.20	1.51	1.30	1.25	1.19

The efficiency is defined as the ratio of the large sample variance of an estimator over that of the ML estimator. The first estimator considered is the IV estimator when x_t and x_{t-1} are used as instruments in (6). It is evident that increasing the number of instruments does not change the asymptotic variance of the IV estimator. The second estimator applied to (6) has been proposed by Cumby et al. (1983), denoted by CHO, with x_t, x_{t-1} and x_{t-2} being the instruments. In this case increasing the number of instruments would lower the asymptotic variance. Irrespective of the number of instruments, in this model, the CHO estimator will not be more efficient than the third estimator that we consider which has been proposed by Hayashi and Sims (1983), HS. This estimator is efficient within the class of GMM estimators based on the orthogonality restrictions in (6) only. Finally in Table 1 we present the relative efficiency of the GLS estimator proposed in Section 2 assuming that $p = 3$. Our numerical results suggest that the GLS estimator based on the more restrictive assumption that $p = 2$ (which coincides with the data generating process) is in this special case fully efficient and therefore the relative efficiency of this estimator is not presented in the table. Probably the efficiency of the GLS estimator with $p = 2$ is due to the fact that the model under consideration is exactly identified.

From Table 1 it is apparent that if $\rho = 0.9$ the differences in relative efficiency between the various estimators can be considerable. In this example the asymptotic variance of the CHO estimator with x_{t-2} included as additional instrument is substantially smaller than that of the standard IV estimator. Moreover it pays either to use more instruments in the computation of the CHO estimator or to use the HS estimator. Finally it is clear from Table 1 that the GLS estimator proposed in Section 2 can in turn be substantially more efficient than the HS estimator.

4. CONCLUSIONS

In this note we demonstrated how the assumption that a future expectation depends on a finite number of variables only can be exploited to increase the efficiency of simple consistent estimators. In the example, the ML estimator appeared to be only about 20 percent more efficient asymptotically than the GLS estimator which is computationally more easy to implement. This finding suggests that in empirical work, it is more appropriate to approximate the unobserved future expectations by a conditional expectation based on a finite number of past observations and then apply GLS than to substitute the future observation and apply some IV method. Note however that the GLS estimator is no longer consistent if the expectation in fact depends on variables in addition to the ones in the auxiliary equation. In this sense the estimator is less robust than GMM procedures. The GLS estimator proposed in this paper can be generalized to related models. Some of these extensions are considered explicitly in Nijman and Palm (1986).

Tilburg University, The Netherlands
University of Limburg, The Netherlands

APPENDIX

Details on the Computation of the GLS Estimator. In order to derive an explicit expression for the GLS estimator in (12) define matrices $V_{\varepsilon\varepsilon}$, $V_{\varepsilon\xi}$ and $V_{\xi\xi}$ as

$$
\begin{aligned}
\text{(A.1)} \qquad V_{\varepsilon\varepsilon}(t,s) &= \sigma_e^2 + 2\rho\lambda\sigma_{eu} + \rho^2\lambda^2\sigma_u^2 && \text{if } t = s\\
&= 0 && \text{if } t \neq s\\
V_{\varepsilon\xi}(t,s) &= \lambda\sigma_{eu} + \rho\lambda^2\sigma_u^2 && \text{if } t = s\\
&= \sigma_{eu} + \rho\lambda\sigma_u^2 && \text{if } t = s+1\\
&= 0 && \text{otherwise,}\\
V_{\xi\xi}(t,s) &= (1+\lambda^2)\sigma_u^2 && \text{if } t = s\\
&= \lambda\sigma_u^2 && \text{if } |t-s| = 1\\
&= 0 && \text{otherwise.}
\end{aligned}
$$

The covariance matrix Ω of η conditional on Z can now be written as

$$
\text{(A.2)} \quad \Omega = V_{\varepsilon\varepsilon} - \rho V_{\varepsilon\xi}Z(Z'Z)^{-1}Z' - \rho Z(Z'Z)^{-1}Z'V'_{\varepsilon\xi} + \rho^2 Z(Z'Z)^{-1}Z'V_{\xi\xi}Z(Z'Z)^{-1}Z'.
$$

As Ω is a $(T \times T)$ matrix, it is computationally attractive not to invert Ω directly but to write

$$
\begin{aligned}
\text{(A.3)} \qquad \Omega &= \Omega_1 + \Omega_2\Omega_3\Omega_2'\\
\Omega_1 &= V_{\varepsilon\varepsilon} - \rho^2 V_{\varepsilon\xi}Z(Z'Z)^{-1}Z'V'_{\varepsilon\xi}\\
\Omega_2 &= Z - \rho^{-1}V_{\varepsilon\xi}Z(Z'V_{\xi\xi}Z)^{-1}Z'Z
\end{aligned}
$$

$$\Omega_3 = \rho^2(Z'Z)^{-1}Z'V_{\xi\xi}Z(Z'Z)^{-1},$$

and to use the matrix inversion lemma

$$\text{(A.4)} \qquad (\Omega_1 + \Omega_2\Omega_3\Omega_2')^{-1} = \Omega_1^{-1} - \Omega_1^{-1}\Omega_2(\Omega_3^{-1} + \Omega_2'\Omega_1^{-1}\Omega_2)^{-1}\Omega_2'\Omega_1^{-1},$$

which requires only inversion of low dimensional matrices if the matrix inversion lemma is applied in a similar way to invert Ω_1. Note that $V_{\varepsilon\xi}Z = (\sigma_{eu} + \rho\lambda\sigma_u^2)(Z_- + \lambda Z)$ which implies that the GLS estimator can be viewed as an instrumental variables estimator with a matrix weighted linear combination of W, Z and Z_- as the matrix of instruments which proves its consistency.

REFERENCES

Cumby, R. E., J. Huizinga and M. Obstfeld, "Two-step Two-stage Least Squares Estimation in Models with Rational Expectations," *Journal of Econometrics* 21 (1983), 333–355.

Hansen, L. P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50 (1982), 1029–1054.

——— and R. J. Hodrick, "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis," *Journal of Political Economy* 88 (1980), 829–853.

Hayashi, F. and C. Sims, "Nearly Efficient Estimation of Time Series Models with Predetermined, But Not Exogenous, Instruments," *Econometrica* 51 (1983), 783–798.

Hoffman, D. L., "Two-step Generalized Least Squares Estimators in Multi-equation Generated Regressor Models," *Review of Economics and Statistics* 59 (1987), 336–346.

McCallum, B. T., "Rational Expectations and the Estimation of Econometric Models: An Alternative Procedure," *International Economic Review* 17 (1979), 484–490.

Nijman, T. E., "Estimation of Models Containing Unobserved Rational Expectations," in F. van der Ploeg, ed., *Advanced Lectures in Quantitative Economics* (London: Academic Press, 1990), 501–531.

——— and F. C. Palm, "Consistent Estimation of Rational Expectations Models," presented at the European Meeting of the Econometric Society in Budapest, 1986.

——— and M. F. J. Steel, "Exclusion Restrictions in Instrumental Variables Equations," Research Memorandum FEW No. 327, Tilburg University, 1988, forthcoming in *Econometric Reviews*.

Pesaran, M. H., *The Limits to Rational Expectations* (New York: Basil Blackwell, 1988).

Wickens, M. R., "The Efficient Estimation of Econometric Models with Rational Expectations," *Review of Economic Studies* 49 (1982), 55–67.